

# (12) UK Patent Application (19) GB (11) 2 347 811 (13) A

(43) Date of A Publication 13.09.2000

(21) Application No 9828720.4

(22) Date of Filing 24.12.1998

(71) Applicant(s)  
Independent Television Commission  
(Incorporated in the United Kingdom)  
33 Foley Street, LONDON, W1P 7LB, United Kingdom

(72) Inventor(s)  
Mohammed Ghanbari  
Kwee Teck Tan

(74) Agent and/or Address for Service  
Boult Wade Tennant  
Verulam Gardens, 70 Gray's Inn Road, London,  
WC1X 8BT, United Kingdom

(51) INT CL<sup>7</sup>  
H04N 17/00

(52) UK CL (Edition R )  
H4F FHHX F3P F3T

(56) Documents Cited  
US 5796875 A

(58) Field of Search  
UK CL (Edition R ) H4F FGXD FHD FHHX  
INT CL<sup>7</sup> G06T 7/00 7/40 , H04N 17/00  
Online: WPI, EPDOC, JAPIO, INSPEC, COMPDX,  
ELSEVIER, Internet

(54) Abstract Title  
**Measuring blockiness in decoded video images**

(57) An apparatus and method for determining the degree of blockiness in decoded video images is disclosed. A Fourier transform of the decoded video image or an image derived from a decoded video image is generated. Components of the Fourier transform characteristic of block edges in the decoded image are identified. The energy of at least one of the identified components is measured, and the measured energy of the or each identified component is compared with the total energy within the Fourier transform. The comparison is indicative of the degree of blockiness in the decoded image prior to generation of the Fourier transform, and can be used to predict an objective degree of blockiness with improved accuracy. By filtering in the frequency, rather than the spatial domain, the block edges within the decoded image are isolated from the remaining information within the decoded image with significantly more accuracy than previously.

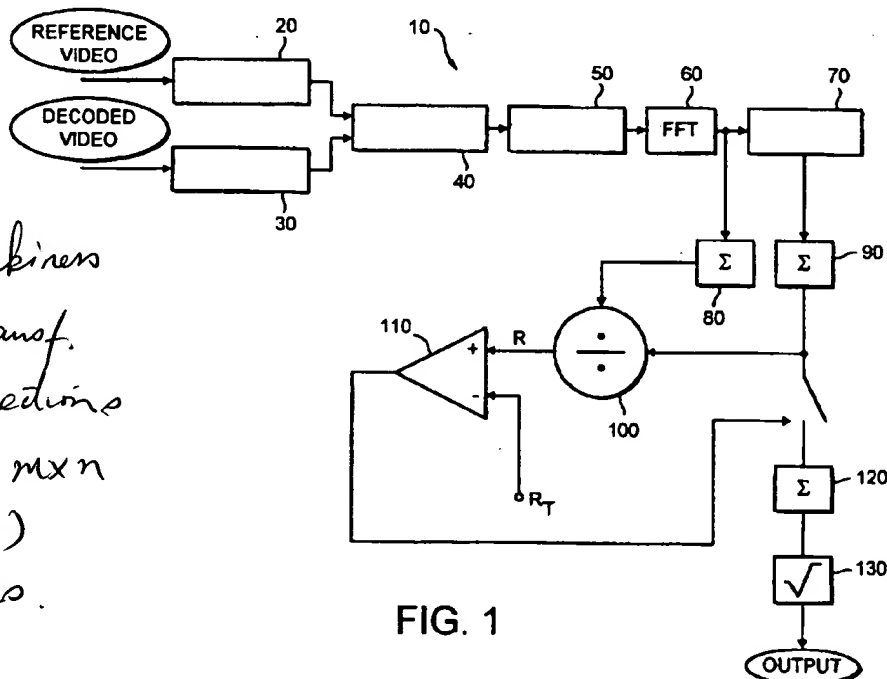


FIG. 1

• Detect blockiness  
by Fourier Transf.  
over sections  
each containing  $m \times n$   
(e.g.  $4 \times 4 = 16$ )  
macroblocks.

GB 2 347 811 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

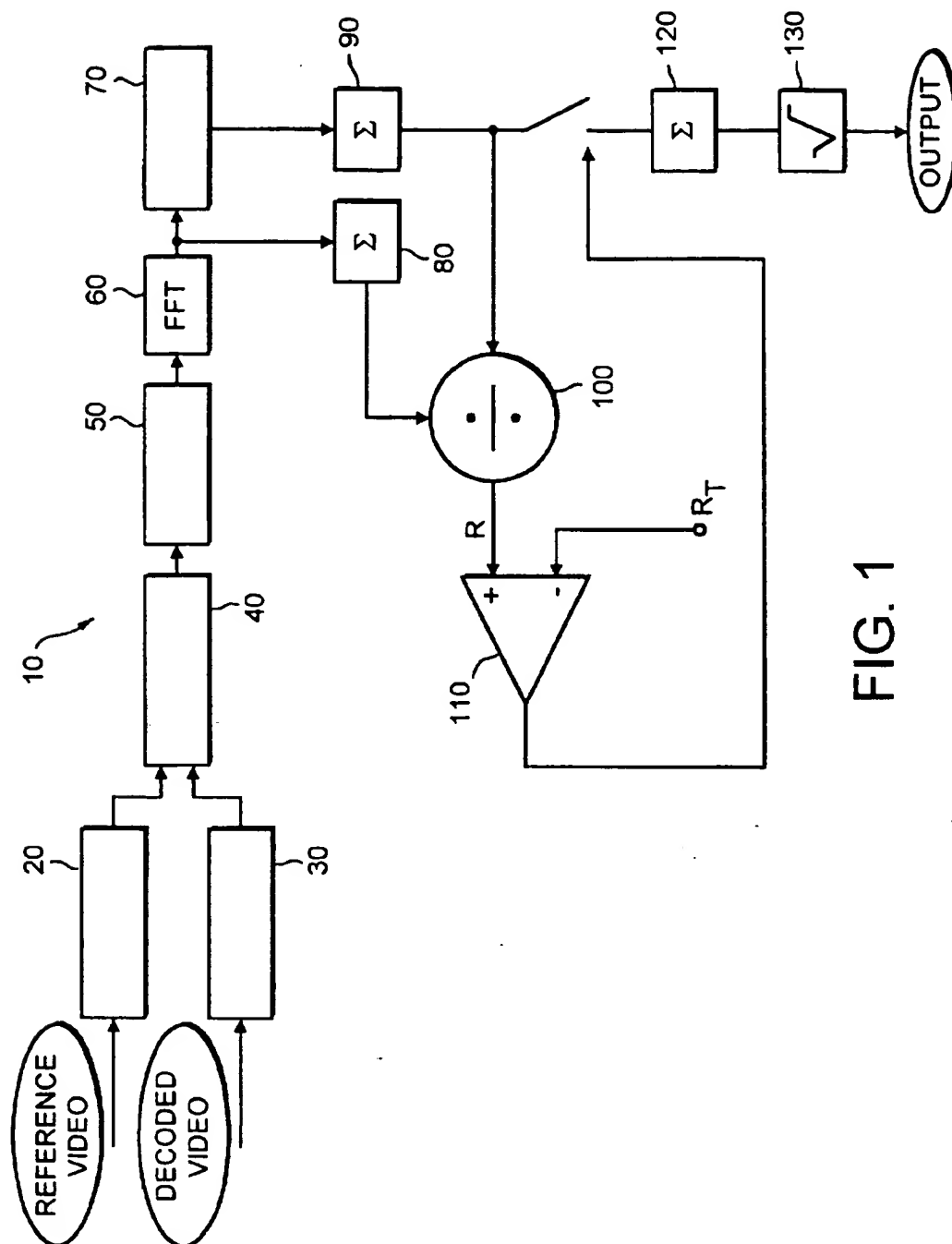


FIG. 1

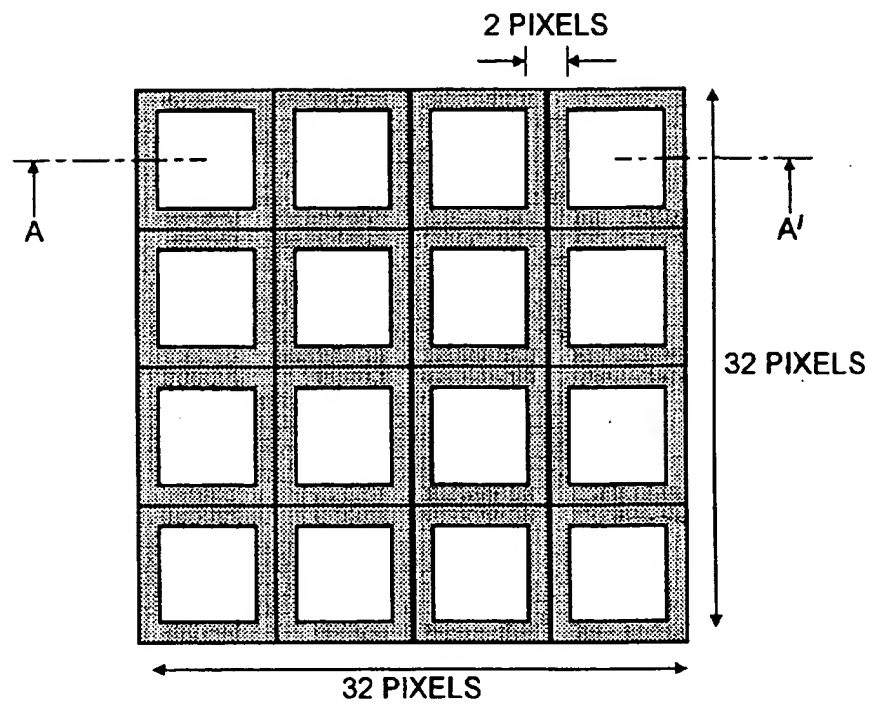


FIG. 2

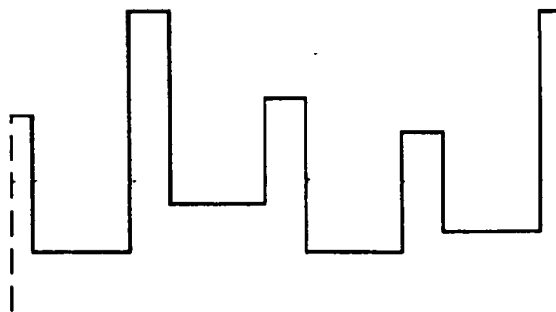


FIG. 3

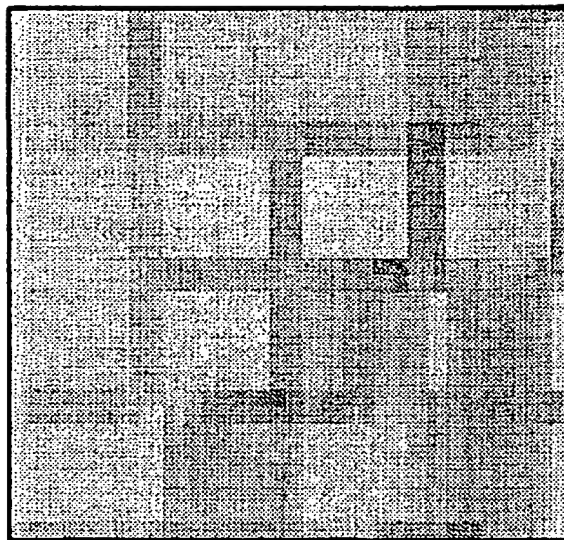


FIG. 4

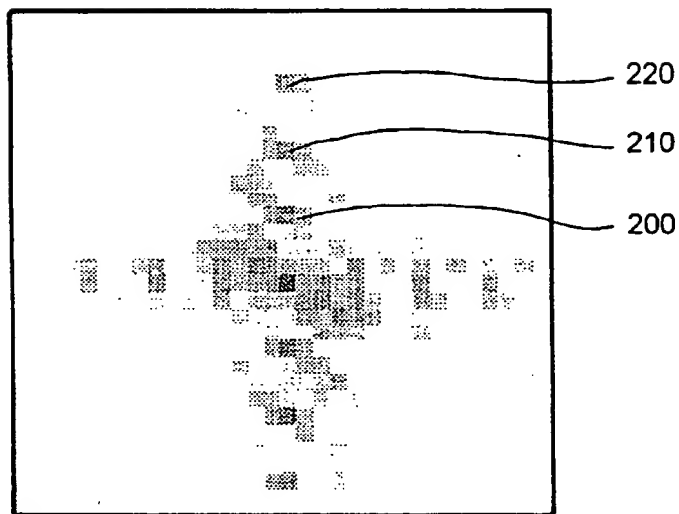


FIG. 5

APPARATUS AND METHOD FOR PROCESSING VIDEO IMAGES

This invention relates to an apparatus and method for processing video images, and in particular for  
5 determining the degree of blockiness in decoded video images.

Transmission of video images tends to reduce their quality. For example, when a modulated analogue video signal is broadcast, losses in the transmission  
10 system cause a degradation in the received signal and a consequent reduction in picture quality.

Transmission of video information in the form of digital data can also result in a received video image of poorer quality than the original image. This  
15 reduction in quality is primarily due to the image compression techniques usually employed to reduce the amount of information that needs to be sent.

One well known system for compressing video data in digital form employs the MPEG compression  
20 algorithm. This algorithm utilizes the fact that there is a significant amount of redundancy in video information. Although a first of a series of consecutive images generally needs to be fully encoded, subsequent images may only change from that  
25 first reference encoded image (called an I frame) by a small amount, as for example in a slow pan shot of a landscape. Frames of video subsequent to the reference image can thus be encoded in more compressed form by encoding the difference between that  
30 subsequent frame and the reference frame in the form of a motion vector. As will be well known to those skilled in the art, frames encoded by reference only to the I-frame are known as P-frames, and frames encoded with reference to both an I-frame and P-frame  
35 (and thus having bi-directional motion vectors) are

known as B-frames.

Any such compression technique will introduce artefacts into images. The more information that needs to be encoded, the more pronounced such artefacts will be. In MPEG coding in particular, the presence of such artefacts as blurring, ringing and blockiness (also known as "tiling") can become quite noticeable to the viewer when the complexity of the video stream increases beyond the capability of the coder/decoder (codec) to encode each picture. For example, motion in the video images, or high textual content, can often cause artefacts in the decoded images.

It is therefore desirable to quantify the quality of both analogue and digital images when decoded. Traditionally, this has been done on a subjective basis by showing both the original image and the received (analogue or digital) image to a panel of viewers and asking them to rate the quality of the received image numerically. Further details of this procedure are described, for example, in "Recommendation ITU-R BT. 500-7 (Revised), 1996, methodology for the subjective assessment of the quality of television pictures".

Whilst this procedure has been successful in measuring the quality of analogue images, where a short series of video images will be generally representative of the quality at any other time in the whole video sequence, the techniques is not so successful for digital images. This is because of the way consecutive images are differently encoded as I, B, or P frames. In any event, using viewers and asking them to score picture quality is both labourious and expensive.

It has been determined that of the different

artefacts introduced by digital video image compression, blockiness is the most visually noticeable. Thus, the ability to quantify the degree of blockiness gives a good indication of the quality of the images seen by the viewers.

A number of electronic (objective) blockiness detectors have been proposed over the years. Two such techniques are described in "A distortion measure for blocking artifacts in images based on human visual sensitivity", IEEE Trans. on Image Processing, VOL. 4, NO. 6, June 1995 pages 713-724, by Karunasekera and Kingsbury, and "Objective measures for detecting digital tiling" T1A1.5/95-104, 1995, by Melcher and Wolf.

Previous techniques suffer a number of drawbacks. The main problem is that the mask used to isolate blocky edges in an image from the actual data content of the image itself tends to be insufficiently accurate. Thus, either some blocky edges are not masked properly, or else straight vertical or horizontal lines which are part of the image and not due to blockiness are incorrectly masked.

It is an object of the present invention to address these problems with the prior art.

According to the present invention, there is provided a method of determining the degree of blockiness in a decoded video image, comprising the steps of generating a Fourier transform of a decoded video image or a Fourier transform of an image derived from the decoded video image; identifying components of the Fourier transform characteristic of block edges in the decoded image; measuring the energy of at least one of the identified components; and comparing the measured energy of the or each identified component with the total energy within the Fourier transform,

the comparison being indicative of the degree of blockiness in the decoded image prior to generation of the Fourier transform.

5 By filtering in the frequency, rather than the spatial domain, the isolation of the blocky edges within the decoded image from the remaining information within the decoded image (which typically corresponds with the information in the image prior to encoding) is substantially more accurate than before.  
10 Previously, a mask was applied in the spatial domain to attempt to separate out the block edges from the remaining information in the video image.

Because the block edges can be better separated in the frequency domain than in the spatial domain,  
15 the degree of blockiness can in turn be predicted, objectively, with significantly better accuracy than previously.

Preferably, the step of generating a Fourier transform of a decoded video image or an image derived  
20 from the decoded video image comprises generating a Fourier transform of a first gradient image derived from the decoded video image. Using the gradient image of the decoded video image, rather than the decoded video image itself, improves the contrast of the  
25 blocky edges in the spatial domain. This improves the detection of the characteristic components in the frequency domain. It will be understood, however, that the application of a Fourier transform to the decoded image itself is equally possible.

30 Preferably, the method further comprises, prior to generation of the Fourier transform, generating an image mask from a gradient image of a corresponding unencoded video image; and applying the image mask to the gradient image of the decoded video image to  
35 selectively enhance the block edges relative to the



remainder of the decoded video image.

Although the majority of the separation of the blocky artefacts from the other types of distortion is mostly achieved in the filtering stage, it is  
5 advantageous to carry out masking in the spatial domain as well. The use of an image stripped of all information other than artefacts (ringing, blurring and blockiness) prevents straight lines, for example, in the unencoded video image (which will generate  
10 characteristic components in the frequency domain), from erroneously being interpreted as blocky edges. Whilst it is preferable to apply the mask to the gradient image of the decoded video image, it is also feasible to apply the mask to the decoded image  
15 itself.

The identified components may comprise the fundamental frequency and at least one harmonic frequency. Moreover, the Fourier transform of the decoded video image may include a first set of  
20 components arising from one or more first block edges in the decoded image, and a second set of components arising from one or more second block edges in the decoded image, the method then further preferably comprising measuring the energy of both the first set  
25 of components and the second set of components. For example, the or each first block edge may be substantially spatially orthogonal with the or each second block edge.

As an illustration, the decoded video image may  
30 have been MPEG encoded. Here, in particular with I-frames, the blockiness is associated with orthogonal vertical and horizontal lines defined by the boundaries of the 8 pixel x 8 pixel DCT blocks. Each raster of the gradient image generated from the  
35 decoded image thus has a luminance which tends to vary

as a rectangular wave of mark-space ratio 1:3. The Fourier transform of this luminance pattern has a fundamental and two harmonic frequencies.

Preferably, the method further comprises  
5 calculating the total energy of the first set of components, and the total energy of the second set of components. This calculation yields the total energy in all of the block walls in a particular image.

Advantageously, the method comprises partitioning  
10 the first gradient image into subgroups of pixels, (each preferably containing a plurality of block edges) Fourier transforms then being separately generated for each of said subgroups of pixels. For example, if MPEG encoding is used, it is preferable to  
15 partition the gradient image of the decoded video image into subgroups of 32 x 32 pixels, each containing sixteen 8 x 8 blocks.

The step of comparing the measured energy of the or each identified component may include determining,  
20 for each subgroup of pixels, a ratio (R) of the energy of those identified components in the Fourier transform to the total energy within the Fourier transform. More particularly, the step of comparing the measured energy of the or each identified  
25 component may include, for each sub-group of pixels, determining a ratio  $R(i, h)$  of the energy of those identified components in the Fourier transform in a first direction thereof, relative to the total energy in that first direction of the Fourier transform; and  
30 determining a ratio  $R(i, v)$  of the energy of those identified components in the Fourier transform in a second direction thereof, relative to the total energy in that second direction of the Fourier transform. In that case, the method may also comprise selecting a  
35 threshold ratio  $R_t$  of both the ratio  $R(i, h)$  and

5  $R(i,v)$ , below which the blockiness of the decoded image is considered to be insignificant, and discarding those subgroups of pixels having both a ratio  $R(i,h)$  and  $R(i,v)$  below the selected threshold ratio  $R_T$ .

10 This procedure has two advantages. Firstly, the total amount of processing of the image is reduced. For example, following the subdivision of the whole image into smaller subgroups of blocks, and deciding (on the basis of the threshold ratio), that no significant blockiness exists in two-thirds of these subgroups, it is possible to ignore these identified blocks in further processing. In other words, only those one third of blocks which exhibit a significant  
15 degree of blockiness need be processed further to give a good estimate of the overall blockiness of the whole decoded video image.

20 A further advantage to subdividing the gradient image into smaller blocks is that localized blockiness information within the decoded video image may be obtained. This can be more useful, in certain situations, than a single blockiness index for the whole decoded image.

25 A preferred way to estimate the overall blockiness is to sum the energy of those subgroups of pixels having a ratio  $R$  above the threshold ratio  $R_T$  to produce a total block edge energy, and then obtain the square root of the total block edge energy.

30 If the encoding of the decoded video image was carried out using the MPEG protocol, then the decoded video image will have been generated from I, P and/or B frames. It is then preferable that the threshold ratio be selected to be different dependent upon whether the decoded video image had been encoded as an  
35 I, P or B frame. This is because the motion vectors in

both P and B frames tend to distort the uniform lines of blockiness which are usually found in I frames. Because B frames include bi-directional motion vectors, the distortion is more pronounced than with P frames. Altering the threshold  $R_t$  depending on the frame type can address this problem. However, it is also possible to use a single threshold  $R_t$  for all frame types, although typically at the expense of a slight reduction in blockiness prediction confidence.

10       The invention also extends to an apparatus for determining the degree of blockiness in a decoded video image, comprising means for generating a gradient image from a decoded video image; means for generating a Fourier transform of the gradient image; 15 means for identifying components of the Fourier transform characteristic of block edges in the decoded image; means for measuring the energy of at least one of the identified components, and means for comparing the measured energy of the or each identified 20 component with the total energy within the Fourier transform, the comparison being indicative of the degree of blockiness in the decoded image prior to generation of the Fourier transform.

25       Further, advantageous, features of this apparatus are set out in the dependent claims appended hereto.

      The invention may be put into practice in a number of ways, one of which will now be described by way of example only and with reference to the figures in which:

30       Figure 1 shows a block diagram of an apparatus for determining the degree of blockiness of a decoded video image according to an embodiment of the present invention;

35       Figure 2 shows an idealized portion of a gradient image obtained from a decoded video image exhibiting

blockiness;

Figure 3 shows the luminance levels of pixels along the line AA' in Figure 2;

Figure 4 shows an actual portion of a gradient image obtained from a decoded video image, again exhibiting blockiness; and

Figure 5 shows a Fourier transform of the portion of the gradient image shown in Figure 4.

Referring first to Figure 1, a schematic block diagram of an apparatus for determining the degree of blockiness in a decoded video image is shown. The apparatus has two inputs, the first of which is a reference video signal, and the second of which is a decoded video signal. The reference video signal includes the original video images - that is, they have at no stage been compressed/coded and then decompressed/decoded. The decoded video signal, on the other hand, is the output of a MPEG codec (not shown) which takes the reference signal as an input, codes it using MPEG coding, and then decodes it for use as a decoded video signal input to the apparatus 10 of Figure 1. It is this decoded video signal which may contain blockiness for the reasons set out above.

The reference video signal and decoded video signal are synchronised with one another. The reference video signal is then fed as an input to a first edge detector 20. The decoded video signal, meanwhile, is fed as an input to a second edge detector 30.

The first edge detector 20 contains a 3 x 3 Sobel filter. The horizontal and vertical luminance gradients at all points in the reference video signal are computed. These orthogonal gradients at each pixel are then combined to generate a single image. This image is known as a gradient image, as will be

familiar to those skilled in the art, and effectively represents a first differential of the luminance as a function of distance.

5 Since only the edges in the gradient image of the decoded video image due to encoding/decoding are of interest, edges due to textural content have to be discarded. This is achieved by masking the gradient image of the decoded picture by a 'mask' generated from the gradient image of a reference image. This second  
10 gradient image is generated in a second edge detector 30, again using a 3 x 3 Sobel filter. The reference image employed corresponds, texturally, to the decoded video image.

15 The mask 40 is a function which decreases with the distance from the point of luminance transition in the second gradient image (derived from the reference video images). Masking is carried out by multiplying the masking function with the second gradient image generated by the second edge detector 30. In  
20 principle, the masking will discard edges in the gradient image of the decoded video image which are due to textural content, without removing edges due to the encoding/decoding process.

25 Following masking in the mask 40 to remove the textural content of the gradient image of the decoded video signal, the masked gradient image is sub-divided at a segmenter 50 into box of 32 x 32 pixels. Thus, typical images are divided into several hundred smaller blocks.

30 MPEG coding employs DCT blocks which are 8 x 8 pixels in size. Thus, each 32 x 32 pixel block produced by the segmenter 50 will appear as an array of 16 tiles, as shown in Figure 2.

35 Figure 3 shows the luminance level of pixels along the line AA' in Figure 2. It will be seen that

the signal approximates to a rectangular signal of mark-space ratio 1:3.

The segmenter 50 time multiplexes each 32 x 32 pixel block generated from the masked decoded gradient image into a processor 60 which carries out a fast Fourier transform on each block centre to it in turn. Because the block edges have a uniform mark-space ratio (Figure 3), the output of the processor 60 contains elements (in the frequency domain) at characteristic frequencies. This feature may be seen by reference to Figures 4 and 5.

Figure 4 shows an actual 32 x 32 pixel block generated by the segmenter 50. Figure 5 shows the 2-D Fourier transform of that block, from which it may be seen that a series of periodic dots on the vertical principal axis are present. The dot labelled 200 constitutes the fundamental component, and the dots 210 and 220 constitute the harmonics. Periodic dots may also be seen on the horizontal principal axis.

The frequency  $f_n$  of the components on the vertical principal axis is defined by

$$f_n = n \times \left( \frac{w}{8} \right), \quad \text{where } n = 1, 2, \dots, \left( \frac{w}{8} - 1 \right)$$

and where  $w$  is the dimension of the block of pixels under analysis (here 32). Thus, in this particular example,  $n = 1, 2$  and  $3$  only, and  $f_1 = 4$ ;  $f_2 = 8$ ; and  $f_3 = 12$ .

Only the components on the horizontal and vertical axes of the Fourier transform calculated by the processor 60 are preserved. The rest are discarded. The Fourier transform is then sent to a first adder 80. It is also sent to a harmonics

extractor 70 which keeps the components at the fundamental and harmonic frequencies but discards the rest. The output of the harmonics extractor 70 is then forwarded to a second adder 90.

5       For each block  $i$  of  $32 \times 32$  pixels, the second adder 90 calculates the total energy of the fundamental and harmonic frequencies in the horizontal direction,  $E(i,h)$  (from the vertical block edges in the spatial domain). The total energy for the vertical  
10       fundamental and harmonic components (from the horizontal block edges in the spatial domain) is also computed to yield  $E(i,v)$ . These two figures are, as will be explained below, a good estimate of the intensity of the blockiness in the decoded video image  
15       that has been input to the apparatus 10. The first adder 80 calculates the total energy in the horizontal and vertical direction separately to give  $E(i,ht)$  and  $E(i,vt)$  respectively.

      For each block  $i$  of  $32 \times 32$  pixels, two ratios,  
20        $R(i,h) = \{E(i,h)/E(i,ht)\}$  and  $R(i,v) = \{E(i,v)/E(i,vt)\}$  are computed in a divider 100. The higher these ratios are, the more likely that there is blockiness in that  $i$ th  $32 \times 32$  pixel block. In order to decide whether a  $32 \times 32$  pixel block contains  
25       blockiness, a threshold ratio  $R_t$  is chosen. Any  $32 \times 32$  pixel blocks having  $R(i,h) < R_t$  and  $R(i,v) < R_t$  are deemed to have a blockiness which would be substantially imperceptible to a human viewer. The comparisons of  $R(i,h)$  and  $R(i,v)$  with  $R_t$  for each  
30       block  $i$  is carried out in a comparator 110.

      Only the  $32 \times 32$  pixel blocks with  $R(i,h)$  greater than or equal to  $R_t$  or  $R(i,v)$  greater than or equal to  $R_t$  proceed to the next stage, a third adder 120. The rest are deemed not to contribute any significant  
35       amount to the final result. In the third adder, the



energy of the fundamental and harmonic frequencies  $E(i,h) + E(i,v)$  for each block with  $R > R_1$  is accumulated. By taking the square root of this accumulated energy, at a square root calculator 130, an output from the apparatus, representing the edge root means square value of the decoded video image input is obtained. This output is chosen as an indicator of the degree of blockiness in the decoded video image.

10        This "blockiness index" may be compared with the degree of blockiness determined by subjective human assessment, to assess its accuracy. As will be familiar to those skilled in the art, there are three different frame types (I, P and B) in a typical MPEG  
15 bit stream. I-frames have fixed block boundaries, so that the edge locations are predictable. For P and B-frames, however, the location of the edges varies, depending upon the motion compensation and the prediction error. The degree of edge variation will  
20 depend upon the magnitude and direction of the motion vectors.

Table 1 shows the correlation between the blockiness index (determined by the apparatus 10 of figure 1), and the blockiness determined subjectively  
25 by a panel of 30 human evaluators using DSCQS methodology (as set out in the above identified Recommendation ITU-R BT.500-7(Revised)). The Pearson correlation in column 3 represents the error between the normalised blockiness index of the apparatus of  
30 Figure 1, and the normalised blockiness index determined by the average score of the 30 human evaluators. The fourth column lists Spearman's Rank Order correlation, which measures the monotonicity.

|    |                    |       |                        |   |
|----|--------------------|-------|------------------------|---|
|    | Picture<br>Type    | $R_T$ | Pearson<br>Correlation | Spearman's<br>Rank Order<br>Correlation |
|    | I frame<br>only    | 0.15  | 0.9286                 | 0.9515                                  |
| 5  | P frame<br>only    | 0.05  | 0.9318                 | 0.9636                                  |
|    | B frame<br>only    | 0.10  | 0.8831                 | 0.8545                                  |
| 10 | I, P & B<br>frames | 0.15  | 0.8388                 | 0.8572                                  |

TABLE 1

It may be seen that the correlation coefficients are high for the various individual picture types, and particularly for I and P-frames. Although P-frames do not have fixed DCT block boundaries, the apparatus of Figure 1 is still well able to predict blockiness. This may be because in P-frames, a group of DCT blocks tend to have identical motion vectors. As a result, the DCT block boundaries are spatially offset by a similar amount, hence maintaining the shape of the lattice seen in Figure 2. Harmonic analysis as described above is relatively insensitive to spacial offset. Thus, the frequency components of the blocky artefacts in the decoded video image are still easily identifiable. Another possibility is that the prediction error is strong enough to smooth out the offset blocky edges due to motion compensation, and impose its own edges around the DCT block boundaries. As a result, the blocky edges remain in the form of a lattice.

For B-frames, on the other hand, the motion compensation is normally effective in making the

prediction error too small to cause blocky edges at the DCT block boundaries. The bi-directional motion compensation also distorts the regular lattice of Figure 2, affecting the ability of the apparatus of Figure 1 to predict blockiness. Nonetheless, the correlation is still high at 0.884 Pearson correlation and 0.85 for Spearman Rank Order correlation.

It will be noted that the threshold ratio chosen for the apparatus differs depending upon the type of frame (I, P or B).  $R_T$  may be determined by maximising the product of Pearson and Spearman's correlation co-efficients between the subjective and objective scores for each group of pictures separately. However, it may be desirable to use only one threshold for all three frame types especially when the picture type is unknown. In this case, the correlation is found to drop to 0.84 (Pearson correlation) and 0.86 (Spearman's correlation).

Although masking of the decoded video image is desirable prior to performing the Fourier transform, it is to be understood that such a procedure is not essential. It may, for example, be important to monitor decoded images only to check for catastrophic blockiness, with precise estimation of the degree of blockiness not being necessary. Such a procedure may be particularly useful in continually monitoring the decoded output of a digital television signal, for example. Carrying out the Fourier transform of the decoded image without first masking on the basis of a reference image means that lines in the decoded image which are meant to be there and are not a result of blockiness will be the subject of the Fourier transform as well. These lines will have components in the frequency domain at characteristic frequencies which may be similar to those components caused by

blockiness, for example if the decoded image includes  
a picture of a tiled floor. Thus, the estimate of the  
degree of blockiness may then be incorrect. However,  
if it is chiefly desired to monitor the general  
5 quality of the decoded image in terms of the amount of  
blockiness, rather than generating an accurate  
objective value indicating the degree of blockiness,  
the avoidance of a mask may be preferable. This is  
because the extra step of masking the decoded video  
10 image (or indeed masking the gradient image thereof)  
increases the overall computational power required,  
and hence the time taken to calculate the degree of  
blockiness.

CLAIMS

1. A method of determining the degree of blockiness in a decoded video image, comprising the steps of:  
5       generating a Fourier transform of a decoded video image or an image derived from the decoded video image;  
          identifying components of the Fourier transform characteristic of block edges in the decoded image;  
10       measuring the energy of at least one of the identified components; and  
          comparing the measured energy of the or each identified component with the total energy within the Fourier transform, the comparison being indicative of  
15       the degree of blockiness in the decoded image prior to generation of the Fourier transform.
2. A method as claimed in claim 1, in which the step of generating a Fourier transform of a decoded video  
20       image or an image derived from the decoded video image comprises generating a Fourier transform of a first gradient image derived from the decoded video image.
3. A method as claimed in claim 2, further  
25       comprising, prior to generation of the Fourier transform;  
          generating an image mask from a second gradient image of a corresponding unencoded video image; and  
          applying the image mask to the first gradient  
30       image of the decoded video image to selectively enhance the block edges relative to the remainder of the decoded video image.
4. A method as claimed in claim 1, 2 or 3, in which  
35       the identified components comprise the fundamental

frequency and at least one harmonic frequency.

5. A method as claimed in any one of the preceding claims, in which the Fourier transform of the video  
5 image or the Fourier transform of the image derived therefrom includes a first set of components arising from one or more first block edges in the decoded image, and a second set of components arising from one or more second block edges in the decoded image, the  
10 method further comprising measuring the energy of both the first set of components and the second set of components.

6. A method as claimed in claim 5, in which the or  
15 each first block edge is substantially spatially orthogonal with the or each second block edge.

7. A method as claimed in claim 5 or claim 6,  
20 further comprising calculating the total energy of the first set of components and the total energy of the second set of components.

8. A method as claimed in any one of claims 2 to 7,  
25 further comprising partitioning the first gradient image of the decoded video image into subgroups of pixels, Fourier transforms then being separately generated for each of the said subgroups of pixels.

9. A method as claimed in claim 8, in which each  
30 subgroup of pixels contains a plurality of block edges.

10. A method as claimed in claim 8 or claim 9 in  
35 which the first gradient image of the decoded video image is partitioned into subgroups of 32 x 32 pixels,

each containing sixteen 8 x 8 blocks.

11. A method as claimed in any one of claims 8, 9 or 10, in which the step of comparing the measured energy  
5 of the or each identified component includes determining, for each subgroup of pixels, a ratio of the energy of those identified components in the Fourier transform to the total energy within the Fourier transform.
12. A method as claimed in claim 11, in which the step of comparing the measured energy of the or each identified component includes determining, for each  
10 sub-group of pixels, a ratio  $R(i, h)$  of the energy of those identified components in the Fourier transform in a first direction thereof, relative to the total energy in that first direction of the Fourier transform; and determining a Ratio  $R(i, v)$  of the  
15 energy of those identified components in the Fourier transform in a second direction thereof, relative to the total energy in that second direction of the Fourier transform.
13. A method as claimed in claim 12, further comprising:  
25 selecting a threshold ratio  $R_t$  of both the ratio  $R(i, h)$  and the ratio  $R(i, v)$ , below which the blockiness of the decoded image is considered to be insignificant; and  
30 discarding those subgroups of pixels having both a ratio  $R(i, h)$  and a ratio  $R(i, v)$  below the selected threshold ratio  $R_t$ .
14. A method as claimed in claim 13, further comprising summing the energy of those subgroups of  
35

pixels having either a ratio  $R(i,h)$  or a ratio  $R(i,v)$  above the threshold ratio  $R_t$ , to produce a total block edge energy, and obtaining the square root of the total block edge energy.

5

15. A method as claimed in any one of the preceding claims, in which the decoded video image has been encoded using the MPEG protocol into I, P and B frames.

10

16. A method as claimed in claim 15 when dependent upon claim 13 or claim 14, in which the threshold ratio is selected to be different dependent upon whether the decoded video image had been encoded as an I, P or B frame.

15

17. A method of determining the degree of blockiness in a decoded video image substantially as herein described with reference to the accompanying figures.

20

18. An apparatus for determining the degree of blockiness in a decoded video image, comprising:  
means for generating a Fourier transform of a decoded video image, or a Fourier transform of an image derived from the decoded video image;  
means for identifying components of the Fourier transform characteristic of block edges in the decoded image;  
means for measuring the energy of at least one of the identified components; and  
means for comparing the measured energy of the or each identified component with the total energy within the Fourier transform, the comparison being indicative of the degree of blockiness in the decoded image prior to generation of the Fourier transform.

25

30

35



19. An apparatus as claimed in claim 18, in which the means for generating a Fourier transform of a decoded video image, or a Fourier transform of an image derived from the decoded video image is arranged to  
5 generate a Fourier transform of a first gradient image of the decoded video image.

20. An apparatus as claimed in claim 19, further comprising:  
10 means for generating an image mask from a second gradient image of a corresponding unencoded video image; and  
means for applying the image mask to the first gradient image of the decoded video image to  
15 selectively enhance the block edges relative to the remainder of the decoded video image.

21. An apparatus as claimed in claim 18, claim 19 or claim 20, in which the identified components comprise  
20 the fundamental frequency and at least one harmonic frequency.

22. An apparatus as claimed in any one of claims 18 to 21, in which the Fourier transform of the decoded  
25 video image, or the Fourier transform of the image derived from the decoded video image, includes a first set of components arising from one or more first block edges in the decoded image, and a second set of components arising from one or more second block edges  
30 in the decoded image, the means for measuring the energy of at least one of the identified components being arranged to measure both the first set of components and the second set of components.

35 23. An apparatus as claimed in claim 22, in which the

or each first block edge is substantially spatially orthogonal with the or reach second block edge.

5       24. An apparatus as claimed in claim 22 or claim 23, further comprising means for calculating the total of the first set of components, and the total energy of the second set of components.

10       25. An apparatus as claimed for any of claims 19 to 24, further comprising means for partitioning the decoded video image into subgroups of pixels, wherein Fourier transforms are separately generated for each subgroup of pixels within the first gradient image of the decoded video image.

15       26. An apparatus as claim 25, in which each subgroup of pixels contains a plurality of block edges.

20       27. An apparatus as claimed in claim 25 or claim 26, in which the decoded video image is partitioned into subgroups of 32 x 32 pixels, each containing sixteen 8 x 8 blocks.

25       28. An apparatus as claimed in any of claims 18 to 27, in which the decoded video image has been encoded using the MPEG protocol into I, P and B frames.

30       29. An apparatus for determining the degree of blockiness in a decoded video image substantially as herein described with reference to the accompanying figures.



INVESTOR IN PEOPLE

Application No: GB 9828720.4  
Claims searched: All

Examiner: Joe McCann  
Date of search: 5 July 2000

## Patents Act 1977 Search Report under Section 17

### Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK CI (Ed.R): H4F(FHHX,FHD,FGXD)

Int CI (Ed.7): H04N(17/00);G06T(7/00,7/40)

Other: Online: WPI, EPODOC, JAPIO, INSPEC, COMPDX, ELSEVIER, Internet.

### Documents considered to be relevant:

| Category | Identity of document and relevant passage | Relevant to claims |
|----------|---|--------------------|
| X        | US 5796875 (READ) - see abstract          | 1,18               |

|   |   |   |  |
|---|---|---|--|
| X | Document indicating lack of novelty or inventive step   | A | Document indicating technological background and/or state of the art.  |
| Y | Document indicating lack of inventive step if combined with one or more other documents of same category. | P | Document published on or after the declared priority date but before the filing date of this invention.          |
| & | Member of the same patent family  | E | Patent document published on or after, but with priority date earlier than, the filing date of this application. |

An Executive Agency of the Department of Trade and Industry

05/27/2004, EAST Version: 1.4.1